



## Dishonesty under scrutiny

Jeroen van de Ven, Marie Claire Villeval

### ► To cite this version:

Jeroen van de Ven, Marie Claire Villeval. Dishonesty under scrutiny. Journal of the Economic Science Association, 2015, 1 (1), pp. 86-99. 10.1007/s40881-015-0002-6 . halshs-01137676

**HAL Id: halshs-01137676**

**<https://shs.hal.science/halshs-01137676>**

Submitted on 3 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Dishonesty under scrutiny

Jeroen van de Ven<sup>a</sup> and Marie Claire Villeval<sup>b</sup>

*This version : 4 February 2015*

**Abstract:** We investigate how different forms of scrutiny affect dishonesty, using Gneezy's (2005) deception game. We add a third player whose interests are aligned with those of the sender. We find that lying behavior is not sensitive to revealing the sender's identity to the observer. The option for observers to communicate with the sender, and the option to reveal the sender's lies to the receiver also do not affect lying behavior. Even more striking, senders whose identity is revealed to their observer do not lie less when their interests are misaligned with those of the observer.

**Keywords:** Deception, lies, dishonesty, social image, experiment.

**JEL Classification:** C91, D83

<sup>a</sup> ACLE (University of Amsterdam) and Tinbergen Institute. Address: Valckeniersstraat 65-67, 1018 XE Amsterdam, The Netherlands. E-mail: [j.vandeven@uva.nl](mailto:j.vandeven@uva.nl).

<sup>b</sup> Université de Lyon. Address: F-69007, France; CNRS, GATE Lyon St Etienne, 93, Chemin des Mouilles, F-69130, Ecully, France; IZA, Bonn, Germany. E-mail: [villeval@gate.cnrs.fr](mailto:villeval@gate.cnrs.fr).

We are grateful to the Editor (Robert Slonim) and two anonymous referees for very valuable comments and suggestions, to B. von Hippel for useful discussions, to R. Cautain for programming the experiment and to Q. Thévenet and Z. Dai for valuable research assistance. This research has been supported by the Research Priority Area of the University of Amsterdam and from a grant of the French National Research Agency (ANR, FELIS grant, ANR-14-CE28-0010-01). It was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

## I. INTRODUCTION

Recently, the economic literature on deception has developed considerably, identifying motivations behind lying (Wang *et al.*, 2010; Gneezy, 2005; Hao and Houser, 2011; Erat and Gneezy, 2012; Ariely, 2012; Pruckner and Sausgruber, 2013; Gibson *et al.*, 2013; Charness *et al.*, 2014). In particular, this literature has identified the importance of both lie aversion (Vanberg, 2008; Lundquist *et al.*, 2009) and guilt aversion (Charness and Dufwenberg, 2006; Battigalli and Dufwenberg, 2007; Ellingsen *et al.*, 2009, 2010), and the importance of the maintenance of the self-concept of honesty (Fischbacher and Föllmi-Heusi, 2013; Mazar *et al.*, 2008; Gino *et al.*, 2009). Some studies report on dishonest behavior when there is some form of monitoring and/or punishment, and therefore a degree of observability (e.g., Mazar *et al.*, 2008; Belot and Schroder, 2013; Reuben and Stephenson, 2013; Greenberg *et al.*, 2014), or in front of an audience (Belot *et al.*, 2009).

In this study, we are interested in situations where the observers have a stake in the outcome. The impact of scrutiny on lying behavior in such a context has remained largely unexplored in the economic literature. We analyze in two laboratory experiments how various conditions of scrutiny affect lying behavior by adding an observer to Gneezy's (2005) sender-receiver deception game. In Experiment 1, both the sender and the observer know the distribution of payoffs in each of two possible states of nature, while the receiver does not. The sender sends a (possibly false) message stating which option yields the highest payoff to the receiver. The receiver must then decide to follow the sender's message or not. If the sender believes that his message will be followed, it is payoff maximizing for the sender to lie ("advantageous lies"). If the sender believes that his message will not be followed, lying is in the receiver's advantage ("disadvantageous lies").

We designed four treatments. In all treatments, the sender's and the observer's monetary payoffs are aligned, and the total payoff of the three participants is kept constant. In the

Baseline treatment, the observer can see the sender's message. This treatment allows us to establish the base rate of lying in the presence of an observer who has no information about the sender's identity and cannot take any action. In the Exposure treatment, the sender's identity is revealed to the observer by letting the sender stand up, which allows us to study if lifting anonymity deters dishonesty. In the Chat treatment, we allow for free communication between the sender and the observer. This allows us to determine whether communication favors the moral norm of honest reporting. Finally, in the Snitch treatment the observer can reveal the sender's lie to the receiver but cannot send a message to the receiver if the sender does not lie. This is made common information to the receiver. Thus, the receiver knows that the sender lied after receiving a message from the observer, but if no message has been received from the observer, she does not know if the sender told the truth or if the sender lied but the observer did not denounce the lie. This treatment enables us to study how the consequence of possible denunciation affects the sender's dishonesty.

Our main results show that in the Baseline, 45.16 percent of the senders send a false message. The lying rate is 26.67 percent when senders expect the message not to be followed, and 62.50 percent when they expect the message to be followed. Surprisingly, these values do not differ significantly between the Baseline and any of the other treatments. The exception is the Snitch treatment, in which senders send more false messages when they expect receivers not to follow (disadvantageous lies). This finding rejects our hypothesis that participants will lie less. We speculate that senders delegate the decision to the observers, which they can only do by sending a false message.

We were especially surprised to find that senders were insensitive to their identity being revealed to observers. One reason could be that some senders are misleading receivers in another way, by strategically telling the truth when they expect the message not to be followed (Sutter, 2009). Another reason could be that they do not care about being identified,

because their monetary incentives are aligned with those of the observer. In Experiment 2 we test the robustness of our findings. We ran two additional Exposure treatments, with ten instead of two possible options. This design (inspired by Erat and Gneezy, 2012) eliminates strategic truthtelling. In one of the treatments, we also introduced a conflict of interest between the sender and the observer. We find that the proportion of advantageous lies is higher than in the original Exposure treatment, but strikingly find that this proportion is the same when interests are aligned and misaligned, showing no evidence of shame or guilt aversion. Overall, our findings suggest that higher scrutiny has little impact on dishonesty when the social cost of lying is limited, and transparency policies are not always a remedy against dishonesty.

## II. EXPERIMENT 1

### *2.1. Design and Treatments*

We use Gneezy's (2005) deception game and add a third player. In each treatment we form groups of three participants: a sender, an observer and a receiver. The game involves two possible payoff allocations. In the *Baseline treatment*, one allocation yields a payoff of €5 to the sender, €6 to the observer and €15 to the receiver. The other allocation yields a payoff of €15 to the sender, €7 to the observer and €4 to the receiver. Note that the sender's and observer's payoffs are aligned. The receiver is unaware of any of the payoffs. The sender and the observer learn which allocation is labeled option X and which one is labeled option Y. The sender has to send one of two possible messages to the receiver: "Option X earns you more than option Y" or "Option Y earns you more than option X". The observer can see the sender's message and this is common knowledge. After receiving the message, the receiver has to choose between options X and Y, knowing that both the sender and the observer know the details of the payoff allocation options. The choice determines which allocation is implemented. Participants played the game only once.

The *Exposure treatment* is identical to the Baseline, except that senders have to stand up one by one before choosing a message. Each observer can identify the sender in his group, but not the other way around. In the *Chat treatment*, the observer and the sender can chat freely with each other during three minutes, using a chat box, before the sender chooses his message. Finally, in the *Snitch treatment*, when the sender lies, the observer can send a message to the receiver that the sender lied to her. Observers cannot send a message to the receiver if the sender does not lie. Hence, observers can only reveal lies, and not send false messages themselves. This way, receivers know that the sender lied after receiving a message from the observer. On the other hand, if the receiver did not receive a message from the observer, she cannot tell if (1) the sender told the truth, or (2) the sender lied but the observer did not denounce the lie.<sup>1</sup> We used the strategy method, asking the observer whether to denounce a lie or not before seeing the sender's message.

Note that the total group payoff is 26 in all the treatments, so that treatment differences cannot be driven by efficiency considerations. In all treatments, the procedures (but not the payoffs or messages) were made common knowledge. We elicited the senders' and observers' beliefs about the proportion of receivers who would follow the message, after the sender sent the message and before the observer observed it. The belief elicitation was not incentivized to avoid the risk of hedging. At the end of the session, we elicited both risk attitudes using a procedure inspired by Gneezy and Potters (1997) and Charness and Gneezy (2012) and the subjects' Machiavellianism score from the MACH-IV test of Christie and Geis (1970).<sup>2</sup>

---

<sup>1</sup> Another possibility would be to allow observers to lie as well but this would make it more complicated for the receiver to interpret any message from the observer. This would also add another strategic element for the sender after telling the truth, as he would have to take into account the observer's strategy and form expectations about whom the receiver will believe after receiving conflicting messages from the sender and observer.

<sup>2</sup> For the risk task, each subject was endowed with 100 tokens (€2) and had to decide how many of these tokens to invest in a risky asset, keeping for himself the amount that is not invested. The investment has a 50 percent probability to be a success. In case of a success, the investment returns 2.5 times the amount invested, and in case of a failure the amount invested is lost. The amount invested is a measure of risk aversion. The responses to the 20 questions of the MACH-IV test, which includes some questions about moral attitudes, generate the Machiavellianism score. A higher score indicates a higher tolerance towards lying as a means to an end.

## 2.2. Procedures

We conducted 29 sessions at GATE-LAB, Lyon, France. The participants were 570 undergraduate students mostly from local engineering and business schools, recruited using ORSEE (Greiner, 2004). Each session had between 12 and 30 participants. Table 1 indicates the characteristics of each treatment. The experiment was computerized using z-Tree (Fischbacher, 2007).

**Table 1. Summary of sessions**

<i>Treatment</i>	<i>Experiment</i>	<i>Nb sessions</i>	<i>Nb participants</i>	<i>%female participants</i>	<i>%female senders</i>
Baseline	I	5	93	49.46	54.84
Chat	I	5	93	54.84	51.61
Snitch	I	5	93	52.69	35.48
Exposure_2	I	5	108	49.07	44.44
Exposure_10	II	4	96	57.29	53.12
Exposure_10_conflict	II	5	87	57.47	58.62
Total	-	29	570	53.33	49.47

Upon arrival, participants were randomly assigned to a terminal and role. Roles were denoted by “A”, “B”, and “C”. Subjects in the role of senders (A) and observers (B) were located in the same room, with all observers seated in the same row and all senders seated in another row (not common knowledge, except in the Exposure treatments). Subjects in the role of receivers (C) were located in an adjacent room, connected through an open door. The instructions were role-specific and not read aloud (see Online Appendix 1 and 2). After answering questions in private, and stressing that the game would only be played once, the experimenters left the room.

It was made common information at the beginning of the experiment that a secretary who was not aware of the content of the experiment would pay participants in cash and in private. Sessions lasted approximately 40 minutes. Average earnings were €11.30 (S.D.=5.13), including the earnings from the risk elicitation task.

## 2.3. Power

With the actual sample size reported in Table 1 we can compute the power of our statistical

tests, based on two-sided tests of proportions for different effect sizes. We assume that the sample size is 31 in each treatment and that the percentage of senders who lie in the Baseline is 52, the percentage found by Gneezy (2005). At the 5 percent significance level, the power of finding a treatment difference of 0.25 (a difference of 25 percentage points), 0.3 and 0.35 is respectively 0.54, 0.72 and 0.87. At the 10 percent significance level, the power of finding a treatment difference of 0.25, 0.3 and 0.35 is respectively 0.67, 0.82 and 0.93. A power of 0.8 is usually considered as acceptable. Thus, roughly, our sample size is large enough to detect treatment differences of 0.3 and higher at the 10% significance level, but not for smaller differences. A treatment difference of 0.3 is fairly large. One should therefore be cautious in interpreting non-significant differences.

#### *2.4. Behavioral conjectures*

Senders choose between a truthful and false message. The receiver's response determines which of the options is implemented. When senders believe that their message will not be followed, lying becomes unattractive. In that case, senders may try to deceive receivers by telling the truth (see Sutter, 2009).

Formally, let  $\pi^h$  and  $\pi^l$  denote the sender's high and low payoff. The number of available options is  $n$ , where one of the options yields  $\pi^h$  and the other  $n - 1$  options yield  $\pi^l$ . When the sender sends a truthful message, or when the observer does not reveal a lie in the Snitch treatment, senders believe that their message is followed with probability  $p$ . In the Snitch treatment, we suppose that senders believe that the observer will reveal a lie with probability  $q$ , with  $q = 0$  by construction in all other treatments. If the observer reveals the lie, it is plausible to assume that the receiver will not follow the sender's message. It is straightforward to show that lying is payoff maximizing if:

$$(1) \quad (\pi^h - \pi^l) \left( p - \frac{1}{n-q} \right) \geq 0.$$

Lying is more attractive when the receiver is more likely to trust the message and the



observer is less likely to reveal a lie. Lying is also more attractive when the number of available options increases. The reason is that only one of the options gives the high payoff to the sender, and if there are many possible options then this particular option is unlikely to be implemented after a truthful message even if the receiver does not trust the message.

We distinguish *advantageous lies* and *disadvantageous lies*. Lies are advantageous if they are payoff maximizing i.e., the sender's beliefs are such that (1) holds.<sup>3</sup> Otherwise, lies are disadvantageous. When we refer to the overall lying rate, we include both types of lies. A *selfish* message is defined as either an advantageous lie or an advantageous truth.

We now formulate our hypotheses. Note that any treatment differences can be caused by either (i) a change in the willingness to lie *conditional* on beliefs (not) satisfying (1), (ii) a change in the proportion of participants with beliefs satisfying (1).

First, we conjecture that the vast majority of senders prefer the allocation (15, 7, 4) to the allocation (5, 6, 15). Given the earlier findings that many people are averse to lying, however, we also expect to see a substantial number of honest messages in the Baseline, even if senders expect that their message will be followed and lying is to their advantage.

*Conjecture 1. In the Baseline treatment, a substantial number of senders send a truthful message.*

In the Exposure treatment, the anonymity of the sender is lifted. Several studies show that people care about their social image (e.g., Benabou and Tirole, 2006; Andreoni and Bernheim, 2009; Ariely *et al.*, 2009; Xiao and Houser, 2011) and we therefore expect that senders are less likely to lie in this treatment because of a possible feeling of shame. We also expect a decrease in disadvantageous lies because observers do not know the senders's beliefs and thus cannot tell if a lie is advantageous or disadvantageous.

---

<sup>3</sup> It is a bit arbitrary to classify lies as advantageous when (1) holds with equality. However, our empirical findings are largely unaffected if we exclude participants for whom this is the case. Note also that the reported beliefs may be biased, which would affect our classification. In particular, beliefs may be self-serving: people may report low beliefs to justify lying. However, observers have no reason to bias their beliefs, and on average senders' beliefs are similar to observers' beliefs in most treatments (see Table 2). The exception is the Baseline, where senders' beliefs are significantly lower than observers' beliefs (two-tailed  $t$ -test,  $p=0.030$ ).

*Conjecture 2: The lying rate is lower in the Exposure treatment compared to the Baseline.*

Conjecture 2 is not trivial. First, the sender's and observer's incentives are aligned, so that observers may actually appreciate lying behavior. Senders may take this into consideration. Charness *et al.* (2007) find that people are more likely to play defect in a Prisoner's Dilemma game when they are observed by group members who also have a stake in the outcome. Second, receivers may also expect a lower lying rate and follow the message more often. This makes lying more attractive. This latter effect is quite subtle, however, and we do not expect it to play an important role. Receivers have very little basis to form their expectations on, and are therefore unlikely to change their behavior. In the analysis, we can control for any changes in beliefs.

The Chat treatment allows senders to motivate their choice and observers to express their moral viewpoint. This may lead to a decrease in the lying rate if observers express disapproval and senders care about their image, as anticipated verbal feedback may increase pro-social behavior (Ellingsen and Johannesson, 2008). Alternatively, this may lead to an increase in the lying rate if the observer shows approval, which may reduce guilt aversion. Senders and observers can also discuss the optimal strategy, changing the sender's beliefs. The effect on lying rates can go either way depending on how beliefs shift. Finally, communication may induce a group identity among the senders and the observers, making the receiver an out-group (e.g., Sutter and Strassmair, 2009), which may increase the lying rate. Considering all possible effects, the effect of communication is ambiguous. We base our conjecture on the findings by Sutter (2009), because the team treatment in that study probably comes closest to our treatment.<sup>4</sup> He finds that lying rates are lower in teams compared to individuals, which is for the most part driven by senders' more pessimistic beliefs. Thus, while the effect can go either way, based on Sutter's findings we formulate our next

---

<sup>4</sup> In his study, teams played the original game by Gneezy (2005) but senders decided as a team which message to send. Team members could interact face-to-face. The possible payoffs were identical for all team members.

conjecture:

*Conjecture 3: The opportunity to chat has a negative effect on the total lying rate due to more pessimistic beliefs by senders that receivers will follow the message. Conditional on beliefs, the lying rate is the same as in the Baseline.*

In the Snitch treatment, observers can tell on senders by revealing the lie to the receiver at a small cost to themselves (they sacrifice €1). Thus, potentially a lie is now observed by both the observer as well as the receiver. This makes lying less attractive for senders who dislike being exposed as dishonest, especially since receivers learn that senders tried to mislead them. Although selfish observers should never denounce, Reuben and Stephenson (2013) found that 32 percent of participants were willing to report the lies of others when groups could not select their members, like in our experiment.<sup>5</sup> Note also that lying becomes less attractive because the message is less likely to be followed (see (1)). The proportion of senders for whom lying is advantageous decreases, which reduces the overall lying rate. There is a possible countereffect to this “snitch effect”. Not getting a message from the observer makes the sender’s message more credible to the receiver. This makes lying more attractive. For the same reason as above, we do not expect receivers to change their behavior substantially, and thus expect that this “credibility effect” is dominated. In the empirical analysis, we control for changes in beliefs and empirically determine the net effect.

*Conjecture 4: In the Snitch treatment, the fear of being denounced decreases the proportion of people for whom lying is advantageous, decreasing the lying rate. Conditional on beliefs, the lying rate is lower than in the Baseline because of the moral cost of denunciation.*

## 2.5. Results

Table 2 reports the overall lying rate in each treatment, lying rates split by advantageous and disadvantageous lies, and percentage of selfish messages (*i.e.* telling the truth when expecting not to be followed or lying when holding the opposite expectation). In line with the previous

---

<sup>5</sup> In their setup, participants can report on others, but their own payoffs are unaffected.

studies (e.g., Gneezy, 2005; Sutter, 2009) we find a substantial proportion of participants who send a truthful message. In the Baseline, 54.84 percent of the senders choose the true message. Even when lying is advantageous, 37.50 percent of senders are honest.

*Result 1. Consistent with conjecture C1, a substantial proportion of individuals tell the truth, even when lying is advantageous.*

**Table 2. Lying rates by treatment**

<i>Treatment</i>	<i>Baseline</i>	<i>Exposure</i>	<i>Chat</i>	<i>Snitch</i>
Lying rate (%)				
All	45.16	47.22	41.93	58.06
Lying is disadvantageous	26.67	25.00	14.29	55.00*
Lying is advantageous	62.50	65.00	64.71	63.64
Selfish message (%)	67.74	69.44	74.19	51.61
Mean reported beliefs that receivers follow (%) <sup>a</sup>				
Senders	47.94	52.97	45.58	54.29
Observers	59.74	50.93	46.19**	59.45
Percentage of receivers following the message <sup>b</sup>	74.19	77.78	74.19	80.00
Mean reported beliefs that observers reveal lie (%)				42.74
Percentage of observers revealing lie				51.61

*Note:* Statistical tests compare treatments to the baseline treatment. \*\* indicate significance at the 0.05 level and \* at the 0.10 level in two-tailed proportion tests (for truths, lies, and following decisions) or Mann-Whitney tests (for mean beliefs). <sup>a</sup> In the Snitch treatment, the value is conditional on the observer not revealing the lie. <sup>b</sup> In the Snitch treatment, the value is conditional on not receiving a message from the observer.

In the Exposure treatment, contrary to conjecture C2, we do not find a decrease in the lying rate. This is true for advantageous as well as disadvantageous lies.

In the Chat treatment, we again find a very similar lying rate compared to the Baseline. The chats do not make senders more pessimistic: average beliefs by senders that their message will be followed are almost the same. Observers, on the other hand, become more pessimistic, lowering their beliefs on average from 59.74 to 46.19 (t-test,  $p=0.017$ ,  $N=62$ ).<sup>6</sup> Observers have more optimistic beliefs than senders in the Baseline, but end up with the same average beliefs in the Chat treatment. Moreover, the beliefs of senders and observers become

<sup>6</sup> Unless specified otherwise, all the non-parametric tests reported in this paper are two-tailed.

highly correlated (Pearson's correlation coefficient 0.498,  $p=0.004$ ,  $N=31$ ), indicating that they discuss their expectations about the receiver's behavior.

If the senders' beliefs and the lying rates are similar across the two treatments, one may suspect that observers did not express much disapproval towards lying. Indeed, upon inspection of the chats, we did not encounter a single conversation in which moral aspects were mentioned, although moral hypocrisy can be detected in one conversation.<sup>7</sup> Participants frequently discussed the fact that receivers may not believe the message (19 out of 29 conversations) and strategies to mislead receivers.<sup>8</sup> Several senders solicit advice from observers (10 out of 29). Those who solicit advice are much less likely to lie: 10 vs. 58 percent (proportion test,  $p=0.013$ ,  $N=29$ ). They also tend to believe that lying is not advantageous: 60 percent of senders who solicit advice have beliefs such that (1) does not hold, vs. 42 percent for senders that do not solicit advice. This difference is not significant, though (proportion test,  $p=0.359$ ,  $N=29$ ).

Turning to the Snitch treatment, we observe that in contrast to conjecture C4, the lying rate is somewhat higher than in the Baseline, although the difference of about 13 percentage points is not significant (proportion test,  $p=0.393$ ). We pointed out before that a higher lying rate is possible if the credibility effect dominates. We find little evidence of this, however. The proportion of people who believe that lying is advantageous (35 percent) is below that in the Baseline (52 percent). There is a remarkable difference with the Baseline, though. Of participants who believe that lying is disadvantageous, 55 percent lie, more than double the percentage of the Baseline (proportion test,  $p=0.094$ ,  $N=35$ ). Possibly, senders prefer to

---

<sup>7</sup> "S(ender): We can either play honest and propose Y, or propose X and the two of us earn more but he will not know. O(observer): By proposing Y, he will perhaps believe that we are lying and choose the other option, which is all benefit for you. That's it. S: Yes, it is what I think. It is very likely that he will choose the other option. O: And then, at least we will have been honest! S: Thus, one can say Y in order to have good conscience while hoping that he will do the opposite! O: Exactly! S: Fine. Then, I'll send him Y."

<sup>8</sup> We also encountered third-level reasoning: "S(ender): We tell the truth and we bet on the fact that she will do the opposite, by choosing message 2. O(bserver): Yes, I think that she will not trust. I agree to tell the truth. S: OK, but if she is ahead of us and thus if she does what we say...we'll be screwed up. O: Yes there is always a risk. But I think it is less likely between the two possibilities."

delegate the decision to the observer, to share responsibility.

An econometric analysis confirms these findings. Table A1 in the Online Appendix reports estimates from a linear probability model. We include the treatments and senders' beliefs about the percentage of receivers who follow the sender's message as explanatory variables. Without any further controls, none of the treatments increases the lying rate significantly compared to the Baseline treatment (column 1). The coefficient of beliefs is positive and significant. When we split the sample by disadvantageous vs. advantageous lies (columns 2 and 3) the lying rate is significantly higher in the Snitch treatment for senders who believe that lying is disadvantageous. Controlling for individual characteristics (gender, age, risk attitude and Machiavellian score) does not affect the estimates (columns 4 and 5).

*Result 2:* Conjectures C2 to C4 are not supported by the data. The lying rate is not lower in the Exposure treatment than in the Baseline. The possibility to chat does not decrease the beliefs about the receivers' trust in the message. Finally the risk of being denounced does not make senders more reluctant to lie; a majority of senders who believe that their message will not be followed are lying as if to delegate the responsibility to the observer.

### III. EXPERIMENT 2

The deception game of Experiment 1 has attractive features, but a disadvantage is that participants may engage in strategic truth telling in order to mislead receivers. Such choices ("strategic truths") cannot be formally classified as lies (although recent neuroeconomic evidence has shown that such strategic truth telling activates similar brain regions as when telling a lie, see Volz *et al.* (2014)). They emanate from the group of participants who believe that lying is disadvantageous to themselves. However, in this group there are also participants that do not wish to deceive receivers per se, but are averse to lying.

To account for this, we ran two additional treatments of the exposure treatment using a design inspired by Erat and Gneezy (2012). The *Exposure\_10* treatment is like the Exposure treatment, except that the number of options and messages is ten instead of two. In only one of those ten options, the sender earns the high payoff; the nine other options give the sender

the low payoff. This largely eliminates incentives for strategic truth telling. Even if senders expect receivers not to follow the message, telling the truth is unlikely to result in the sender's preferred option. In *Exposure\_10\_Conflict*, we introduced a conflict of interest between the sender and the observer. It is similar to the *Exposure\_10* treatment, except that the possible payoff vectors are (15, 7, 4) and (5, 15, 6). Note that the total earnings are still 26 in both cases. Wiltermuth (2011) and Gino *et al.* (2013) have shown that the mere presence of a third-party who benefits from a dishonest behavior allows individuals to self-justify unethical behavior that increases their own payoff. This justification disappears when their payoffs are misaligned. Furthermore, Greenberg *et al.* (2014) show that the lying rate is reduced if the receiver is informed about whether or not the sender has lied. Our conjecture is that the misalignment of payoffs reduces lying behavior.

Table 3 reports the lying rates for those two treatments. We focus on participants who believe that lying is advantageous, which is the vast majority of participants. In *Exposure\_10*, the lying rate for this group is 83, more than 20 percentage points higher than in the Baseline.<sup>9</sup> Thus, even when there can be little doubt about their intentions, senders are still very willing to lie when they can be identified. Strikingly, even in *Exposure\_10\_conflict*, where the payoffs are misaligned, the lying rate is not reduced.<sup>10</sup> The difference in payoffs between lying successfully and reporting the truth is so large that the temptation of increasing one's payoff apparently exceeds the moral cost of dishonesty, even when dishonesty is publicly exposed.

*Result 3. When strategic truths are ruled out by design, the temptation of an advantageous lie is stronger than the moral cost associated with the public display of dishonesty even in front of one's potential victim and independent on whether or not there is a conflict of interest.*

---

<sup>9</sup> The difference with the Baseline may be due to the fact that in nine out of ten options, the sender earns less than the receiver, which he may feel as an excessively disadvantageous situation.

<sup>10</sup> The percentage of selfish messages is lower in *Exposure\_10\_conflict* because the number of participants with beliefs below the threshold is higher (5 instead of 2). The difference is not significant (two-sided proportion test,  $p=0.273$ ).

**Table 3. Summary statistics - Ten-option treatments**

<i>Treatment</i>	<i>Exposure_10</i>	<i>Exposure_10_conflict</i>
Lying rate (%)		
All	84.38	82.76
Lying is disadvantageous	100.00	100.00
Lying is advantageous	83.33	79.17
Selfish message	78.13	65.52
Mean reported beliefs that receivers follow (%)		
Senders	55.84	40.38**
Observers	42.97	48.21
Percentage of receivers following	62.50	75.86

*Note:* Statistical tests compare treatments to Exposure\_10. \*\* indicate significance at the 0.05 level in two-tailed proportion tests (for truths, lies, and following decisions) or Mann-Whitney tests (for mean beliefs).

#### IV. CONCLUSION

Our experiments show that scrutiny has little effect on dishonest behavior. Being identified by another person does not reduce lying, even when lying is unambiguously advantageous and even when there is a conflict of interest with the observer. Our participants seem to feel neither shame nor guilt when lying. This is consistent with behavior observed in real settings when deception permeates organizations.

If some other studies do find an effect of making actions or identities observable to others, why don't we? A possibility is that our study is underpowered. It is not the case, however, that we find a substantial treatment effect but with standard deviations that are too high to be able to reject the null hypothesis. Rather, the differences in lying rates between treatments tend to be rather small.<sup>11</sup> Assuming that there really is no treatment effect, we now offer some alternative explanations for our results. First, one plausible explanation is that many people regard lying as justified in our set up and this affects the senders' normative beliefs. For instance, in Experiment 1, the sender can try to help the receiver, but then the

<sup>11</sup> One exception is when we focus on senders in Experiment I who believe that less than 35 percent of receivers will follow the message. For that subsample, we find that the lying rate is roughly 25 percent lower in the exposure and chat treatments than in the baseline. The sample is, however, rather small; only 34 senders have such beliefs over all four treatments.



sender is hardly any better off than the receiver would be if the receiver were to be deceived. Inequality remains large. Senders may anticipate that observers will regard lying as justified because inequality can be resolved in no way and the receiver cannot morally suffer from being deceived since he will never know about it. In this context an advantage to the first mover can be tolerated by the observer. This effect could be even reinforced when communication is made possible between the sender and the observer who can form an alliance against the receiver. Therefore, due to the impact of the high cost of being honest on the senders' normative expectations, senders may not mind being exposed as liars. This could also explain the difference between our results and those of Greenberg *et al.* (2014) who find a lower lying rate if receivers are told whether or not the sender lied to them. Receivers are not aware of the possible payoffs, and therefore cannot tell if the lie was justified or not. Thus, in their context, senders may be more likely to experience feelings of shame than senders in our experiment.

A second explanation could be that, contrary to a dictator game, the sender does not have full control over the outcome, as it is the receiver that makes the ultimate decision. In such cases, the sender's intentions (e.g., to deceive someone else) are weighted as less important (see Pan and Xiao, 2014). This probably reduces the sender's feeling of responsibility. The impact of scrutiny could possibly be higher if the receiver was passive. Third, in Experiment 1 a deceptive sender may keep a good image of himself because his beliefs are not made visible to the observer; he could always pretend –even to himself– that he lies to help a receiver who he thinks would probably not follow his message. Fourth, in our set up senders are forced to send a message, and even the true message may mislead the receiver. Lies may be condemned more by the observers if there was also an option to abstain from sending a message. These dimensions may be so strong that the lying rate does not fall when the sender lies even in the face of his potential victim. Fifth, perhaps people care about being observed

*per se*, no matter if they remain anonymous or not. Future work may include a treatment in which only the sender is informed about all possible payoffs.

Our findings imply that a transparency policy is not sufficient to deter lying. This suggests exploring other policies, like the punishment of dishonesty or rewarding honesty, or the development of codes of ethics. For example in a three-player game similar to ours, Xiao (2013) has shown that punishment of the sender by the observer can reduce lying behavior but only when the punisher cannot profit himself from the punishment. Of course, several extensions can be thought of. In particular, it would be interesting to measure the impact of priming codes of conduct, or vary the payoff distribution to introduce more equal distributions either in case of honesty or in case of dishonesty.

## References

- Andreoni, J., and B.D. Bernheim. 2009. "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica* 77(5): 1607–1636.
- Ariely, D., A. Bracha, and S. Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99(1), 544–555.
- Ariely, D. 2012. *The (Honest) Truth About Dishonesty*. New-York: Harper Collins Publishers.
- Battigalli, P., and M. Dufwenberg. 2007. "Guilt in Games." *American Economic Review* 97: 170–176
- Benabou, R., and J. Tirole, 2006. "Incentives and prosocial behavior." *American Economic Review* 96(5): 1652–1678.
- Belot, M., V. Bhaskar, and J. van de Ven, 2009. "Promises and Cooperation: Evidence from a TV Game Show." *Journal of Economic Behavior & Organization* 73(3): 396–405.
- Belot, M., and M. Schroder, 2013. "Does Monitoring Work? A Field Experiment with Multiple Forms of Counterproductive Behaviour." Mimeo.
- Charness, G., and M. Dufwenberg. 2006. "Promises and partnership." *Econometrica* 74(6):1579–1601.
- Charness, G., and U. Gneezy. 2012. "Strong evidence for gender differences in risk-taking." *Journal of Economic Behavior & Organization* 83: 50–58.
- Charness, G., D. Masclet, and M.C. Villeval. 2014. "The dark side of competition for status." *Management Science* 60(1): 38–55.
- Charness, G., L. Rigotti, and A. Rustichini. 2007. "Individual Behavior and Group Membership." *American Economic Review* 97(4): 1340–1352.
- Christie, R. and F. Geis. 1970. *Studies in Machiavellianism*. New-York: Academic Press.
- Ellingsen, T. and M. Johannesson. 2008. "Anticipated verbal feedback induces altruistic behavior." *Evolution and Human Behaviour* 29(2), 100–105.
- Ellingsen, T., M. Johanneson, and J. Lilja. 2009. "Trust and Truth." *Economic Journal* 119 (534): 252–276.
- Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik. 2010. "Testing guilt aversion." *Games and Economic Behavior* 68: 95–107.
- Erat, S., and U. Gneezy. 2012. "White Lies." *Management Science* 58(4): 723–733.
- Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10(2): 171–178.
- Fischbacher, U., and F. Föllmi-Heusi. 2013. "Lies in Disguise. An experimental study on cheating." *Journal of the European Economic Association* 11(3): 525–547.
- Gibson, R., C. Tanner, and A. Wagner. 2013. "Preferences for truthfulness: Heterogeneity among and within individuals." *American Economic Review* 103(1): 532–548.
- Gino, F., S. Ayal, and D. Ariely. 2009. "Contagion and differentiation in unethical behavior: the effect of one bad apple on the barrel." *Psychological Science* 20(3): 393–398.
- Gino, F., S. Ayal, and D. Ariely. 2013. "Self-serving altruism? The lure of unethical actions that benefit others." *Journal of Economic Behavior & Organization* 93: 285–292.

- Gneezy, U. 2005. "Deception: The role of consequences." *American Economic Review* 95 (1): 384–394.
- Gneezy, U., and J. Potters. 1997. "An experiment on risk taking and evaluation periods." *The Quarterly Journal of Economics* 112(2): 631–645.
- Greenberg, A., P. Smeets and L. Zhurakhovska (2014). "Lying, Guilt, and Shame." Mimeo.
- Greiner, B. 2004. "An Online Recruitment System for Economic Experiments." Kurt Kremer, Voler Macho (eds). *Forschung und wissenschaftliches Rechnen, GWDG Bericht 63*, Gottingen: Ges. fr Wiss. Datenverarbeitung, 79–93.
- Hao, L., and D. Houser, 2011. "Honest lies." Discussion Paper, ICES, George Mason University.
- Holt, C., and S. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92 (5): 1644–1655.
- Lundquist, T., T. Ellingsen, E. Gribbe, and M. Johannesson. 2009. "The aversion to lying." *Journal of Economic Behavior & Organization* 70: 81–92.
- Mazar, N., O. Amir, and D. Ariely. 2008. "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." *Journal of Marketing Research* 45 (6): 633–644.
- Murphy, K.R. 1993. *Honesty in the workplace. The Cypress series in work and science*. Belmont, CA: Thomson Brooks/Cole publishing Co.
- Pan, X and E. Xiao (2014). "It's Not Just the Thought that Counts: An Experimental Study on Hidden Cost of Giving." Mimeo. Available at SSRN: <http://ssrn.com/abstract=2523753>.
- Pruckner, G.J., and R. Sausgruber. 2013. "Honesty on the Streets: A Field Study on Newspaper Purchasing." *Journal of the European Economic Association* 11(3): 661–679.
- Reuben, E., and M. Stephenson. 2013. "Nobody likes a rat: on the Willingness to report lies and the consequences thereof." *Journal of Economic Behavior & Organization* 93: 384–391.
- Sutter, M. 2009. "Deception Through Telling the Truth?! Experimental Evidence From Individuals and Teams." *The Economic Journal* 119: 47–60.
- Sutter, M. and C. Strassmair (2009), Communication, cooperation and collusion in team tournaments - An experimental study, *Games and Economic Behavior* 66 (1): 506–525.
- Vanberg, C. 2008. "Why do people keep their promises? An experimental test of two explanations." *Econometrica* 76 (6): 1467–1480.
- Volz, K., K. Vogeley, M. Tittgemeyer, Y. von Cramon, and M. Sutter, M. (2014). "The neural basis of deception in strategic interaction." Mimeo.
- Wang, J.T., M. Spezio, and C.F. Camerer. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games." *American Economic Review* 100(3): 984–1007.
- Wiltermuth, S.S. 2011. "Cheating more when the spoils are split." *Organizational Behavior and Human Decision Processes*, 115(2): 157–168.
- Xiao, E., and D. Houser. 2011. "Punish in public." *Journal of Public Economics* 95(7-8): 1006–1017.
- Xiao (2013). "Profit-seeking punishment corrupts norm obedience." *Games and Economic Behavior* 77:321–344.

## Online Appendix 1. Instructions for the two-option treatments (translated from French)

*These are the instructions for all the two-option treatments. The parts of the instructions that are specific to the Chat, Snitch and Exposure\_2 treatments are indicated in brackets.*

### A) Instructions for the senders

Thank you for participating in this short experimental session on decision-making. This session consists of two completely independent parts. We have distributed the instructions for the first part. You will receive the instructions for the last part later.

You will be paid individually in cash in a separate room by someone who is not aware of the content of the experiment.

For the duration of the experiment, your decisions are anonymous; you will never enter your name into the computer. Please remain silent throughout the experiment. These instructions are for your own use only.

### Part 1

You are called "**person A**" and you will be randomly matched with two other people who participate in this session, called "person B" and "person C."

None of you will ever know the identity of the other two participants.

*[In the Exposure\_2 treatment, the previous sentence was replaced with the following paragraph: You will never know the identity of the other two participants. Person C with whom you are matched will never know your identity or that of person B. In contrast, person B with whom you are matched will be able to identify you in the room. Indeed, at the beginning of this part before you make your decision, we will ask each person A to stand up, one by one. We will call you by the number on your tag and you will have to stand up (for example, if you have received the tag with number 1, it means that you are person A number 1). Person B with whom you are matched knows the number of your tag so that this person will be able to identify you in the room. In contrast, you will not be able to identify the person B that you are matched with. We asked person B not to make any sign that would allow you to identify him in the room. Person C knows that person B has identified you and that you cannot identify person B.]*

Two possible payment options are available to you and your co-participants in this part. The two payment options are:

€ 5 for you, 6 € for person B and 15 € for person C

or:

€ 15 for you, 7 € for person B and 4 € for person C.

At the beginning of this part, you and person B (but not person C) learn which distribution is called "option X" and which is called "option Y."

Person C must choose between option X and option Y to determine which option will be implemented. The only information that person C possesses is the information that you send him in a message *[Snitch treatment: and possibly the information that person B sends to him]*. In other words, person C will never know the monetary payoffs associated with each option, nor the differences between the earnings. Person C will never be informed about which option is best for each of the three persons.

We will ask you to choose one of the two following messages to send to person C:

Message 1: "Option X earns you more money than option Y."

Message 2: "Option Y earns you more money than option X."

We will send your message to person C and ask him to choose between option X and option Y. The choice of person C will determine the earnings of the three persons. However, person C will never know the amounts that were offered in the option that was not selected (that is to say, he will never know if your message was true or not).<sup>12</sup> In addition, person C will never know the amounts that you and person B earn depending on the different

---

<sup>12</sup> The sentence: "that is to say, he will never know if your message was true or not" was not displayed in the

options.

Person B receives the same information as you do: B is informed about the payoffs associated with each option, X and Y, and knows the two messages that you can send to person C. Person B is informed about the message you decide to send to person C, but he cannot send any message himself to person C. Person B makes no decision.<sup>13</sup> Both person B and person C know that you know the payoffs associated with each option.

*[Chat treatment only: Before you make your decision which message to send to person C, you have the opportunity to interact with person B using a chat box that appears on your screen. The only restriction regarding your interaction with person B is that it is forbidden to identify yourself or ask person B for identification, use foul language or threats. You have 3 minutes to chat with person B. Person C knows that you can chat with person B but he will never know the content of your discussion. Once the 3 minutes have elapsed, only you will enter on your computer the message that you have decided to send to person C. ]*

*[Snitch treatment only: Before seeing the message that you have selected, person B must decide whether or not to send the following message to person C in case your message is not true: The message from person A is not true. Option ... earns you more money than option .... ]*

*- If your message is not true, the message from person B will be sent to person C only if B has decided to send a message.*

*- If your message is true, no message from person B is sent to person C.*

*Person B will then be informed about the message that you actually chose to send to person C. But you will not know whether person B has decided to send a message to person C if your message is not true. Person C knows that if he receives a message from person B, it means that your message is not true. If he does not receive a message from person B, he does not know whether this is because your message is true or because person B has decided not to send him a message.]*

Person B will be informed about the message that you send to person C but he cannot send any message to person C.<sup>14</sup>

The following table summarizes the information possessed by persons A (you), B and C:

Person	Information
You	<p>The existence of persons B and C</p> <p><i>[Exposure_2 treatment only: The fact that person B with whom you are matched has identified you in the room]</i></p> <p>The payoffs of each person for each option</p> <p>Which option is denoted "X" and which is denoted "Y"</p> <p>The two possible messages</p> <p><i>[Chat treatment only: The possibility to chat with person B]</i></p> <p><i>[Snitch treatment only: The possibility for person B to send the other message to person C if your message is not true. Note: If your message is not true, you do not know whether or not person B has sent a message to person C.]</i></p> <p>Person B learns the message you decided to send to person C.</p> <p>Your payoff after person C has chosen an option</p>
What person B knows	<p>The existence of persons A and C</p> <p><i>[Exposure_2 treatment only: Person B with whom you are matched has identified you in the room]</i></p> <p>The earnings of each person for each option</p> <p>Which option is denoted "X" and which is denoted "Y"</p> <p>The two possible messages</p> <p><i>[Chat treatment only: The possibility to chat with you]</i></p> <p><i>[Snitch treatment only: The possibility to send the other message to person C only if your message is not true]</i></p>

Snitch treatment.

<sup>13</sup> The sentences: "but he cannot send any message himself to person C. Person B makes no decision" were not displayed in the Snitch treatment.

<sup>14</sup> This sentence "but he cannot send any message to person C" was not displayed in the Snitch treatment.

	<p><i>The fact that you (person A) does not know whether or not he sent a message to person C.]</i></p> <p>The message you (person A) decided to send to person C</p> <p>His payoff after person C has chosen an option</p>
What person C knows	<p>The existence of persons A and B</p> <p><i>[Exposure_2 treatment only: Person B with whom you are matched has identified you in the room but you cannot identify who is person B]</i></p> <p>The two possible messages</p> <p><i>[Chat treatment only: The chat opportunity between you and person B]</i></p> <p>The message that you (person A) decided to send to him</p> <p><i>[Snitch treatment only: The fact that receiving a message from person B means that your message is not true]</i></p> <p><i>The fact that you do not know whether or not he received a message from person B.]</i></p> <p>Person B learns the message that you (person A) decided to send to person C.</p> <p>His payoff after choosing an option</p> <p>Note: C knows neither the payoffs at stake, nor the payoffs associated with each option, nor which option is the best for each of the three persons</p> <p><i>[Chat treatment only: nor the content of the discussion between you and person B.]</i></p>

All decisions are taken only once.

At the end of the session, you, person B and person C will only be informed about one's own payoff. In the second part, you will no longer be matched with any participant.

Please, read these instructions again. If you have any question, raise your hand and we will answer your questions in private.

## Part 2

*[The instructions for Part 2 were displayed on the subjects' screen after completion of part 1.]*

### Description of the task

You receive 100 tokens (with 100 tokens= €2). We ask you to choose the number of tokens (between 0 and 100, 0 and 100 included) that you are willing to invest in a risky investment. You keep the tokens that you do not invest.

### The investment

The investment is successful with chance one out of two. If it is a success, you receive 2.5 times the number of tokens you have invested. If the investment is not a success, you lose the tokens you have invested.

Example 1. You invest 0 tokens. You earn  $(100-0)=100$ .

Example 2. You invest 50 tokens. If the investment is a success, you earn:  $(100 - 50) + 2.5*50= 175$ . If the investment is not a success, you earn:  $(100-50) + 0 = 50$ .

Example 3. You invest 100 tokens. If the investment is a success, you earn:  $(100 - 100) + 2.5*100= 250$ . If the investment is not a success, you earn:  $(100-100) + 0 = 0$ .

### How do we determine if your investment is a success?

We ask you to choose a color among the two following colors: red and black. Then, click the button "random draw". Each color has a chance of one out of two to be randomly drawn by the computer program.

If the randomly drawn color is the color that you have selected, the investment is a success.

If the randomly drawn color is not the color you have selected, the investment is not a success.

To sum up: you choose the number of tokens to invest, then you choose the color. Then you click the button for the random draw.

Your earnings will be added to your previous earnings.

-----

## B) Instructions for the observers

Thank you for participating in this short experimental session on decision-making. This session consists of two completely independent parts. We have distributed the instructions for the first part. You will receive the instructions for the last part later.

You will be paid individually in cash in a separate room by someone who is not aware of the content of the experiment.

For the duration of the experiment, your decisions are anonymous; you will never enter your name into the computer. Please remain silent throughout the experiment. These instructions are for your own use only.

### Part 1

You are called "**person B**" and you will be randomly matched with two other people who participate in this session, called "person A" and "person C."

None of you will ever know the identity of the other two participants.

*[In the Exposure\_2 treatment, the previous sentence was replaced with the following paragraph: Persons A and C with whom you are matched will never know your identity. You will never know the identity of person C. In contrast, you will be able to identify the person A with whom you are matched in the room. Indeed, at the beginning of this part before decisions are made, we will ask each person A to stand up, one by one. We will call each person A by the number on his tag and person A will have to stand up. You know the number of the tag of the person A with whom you have been matched: it is the number indicated on the tag we gave you (for example, if you have received the tag with number 1, it means that you are matched with person A number 1). Thus, you will be able to identify the person A with whom you are matched. In contrast, person A will not be able to identify you. We ask you not to make any sign that would allow person A to identify you in the room. Person C knows that you can identify person A and that person A cannot identify you.]*

Two possible payment options are available to you and your co-participants in this part. The two payment options are:

€ 5 for person A, 6 € for you and 15 € for person C

or:

€ 15 for person A, 7 € for you and 4 € for person C.

At the beginning of this part, you and person A (but not person C) learn which distribution is called "option X" and which is called "option Y."

Person C must choose between option X and option Y to determine which option will be implemented. The only information person C possesses is the information that person A sends him in a message [*Snitch treatment: and possibly the information that you send to him*]. Person C will never know the monetary payoffs associated with each option, nor the differences between the earnings. Person C will never be informed about which option is best for each of the three persons.

We will ask person A to choose one of the two following messages to send to person C:

Message 1: "Option X earns you more money than option Y."

Message 2: "Option Y earns you more money than option X."

We will send the message by person A to person C and ask him to choose between option X and option Y. The choice of person C will determine the earnings of the three persons. However, person C will never know the amounts that were offered in the option that was not selected (that is to say, he will never know if the message of person A was true or not).<sup>15</sup> In addition, person C will never know the amounts that you and person A earn depending on the different options.

---

<sup>15</sup> The sentence: "that is to say, he will never know if your message was true or not" was not displayed in the Snitch treatment.



*[Chat treatment only: Before person A makes his decision about the message to send to person C, you have the opportunity to interact with person A using a chat box that appears on your screen. The only restriction regarding your interaction with person A is that it is forbidden to identify yourself or to ask person A for identification, use foul language or threats. You have 3 minutes to chat with person A. Person C knows that you can chat with person A but he will never know the content of your discussion. Once the 3 minutes have elapsed, only person A will enter on his computer the message that he has decided to send to person C. ]*

*[Snitch treatment only: Before seeing the message that person A has selected, you must decide whether or not to send the following message to person C in the case the message sent by person A is not true: The message from person A is not true. Option ... earns you more money than option .... ]*

*- If person A's message is not true, your message will be sent to person C only if you have decided to send a message.*

*- If person A's message is true, you cannot send a message to person C.*

*You will then be informed of the message that person A actually chose to send to person C. But person A will not know whether you have decided to send a message to person C if his message is not true. Person C knows that if he receives a message from you, it means that the message of person A is not true. If he does not receive a message from you, he does not know whether this is because person A's message is true or because you have decided not to send him a message.]*

You are informed about the message that person A decided to send to person C. You have no decision to make.<sup>16</sup> Both person A and person C know that: you know the payoffs associated with each option, the two possible messages sent by A, the message chosen by A; you do not make a decision.<sup>17</sup>

The following table summarizes the information possessed by persons A, B (you), and C:

Person	Information
What person A knows	<p>The existence of persons B and C</p> <p><i>[Exposure_2 treatment only: The fact that you have identified him in the room]</i></p> <p>The payoffs of each person with each option</p> <p>Which option is denoted "X" and which is denoted "Y"</p> <p>The two possible messages</p> <p><i>[Chat treatment only: The possibility to chat with you]</i></p> <p><i>[Snitch treatment only: The possibility for you to send another message to person C if person A's message is not true. Note: If his message is not true, person A does not know whether or not you have sent a message to person C.]</i></p> <p>You (person B) learn the message he decided to send to person C.</p> <p>His payoff after person C has chosen an option</p>
You	<p>The existence of persons A and C</p> <p><i>[Exposure_2 treatment only: You can identify person A with whom you are matched in the room]</i></p> <p>The earnings of each person for each option</p> <p>Which option is denoted "X" and which is denoted "Y"</p> <p>The two possible messages</p> <p><i>[Chat treatment only: The possibility to chat with person A]</i></p> <p><i>[Snitch treatment only: The possibility to send the other message to person C only if person A's message is not true]</i></p> <p><i>The fact that person A does not know whether or not you sent a message to person C.]</i></p> <p>The message person A decided to send to person C</p> <p>Your payoff after person C has chosen an option</p>
What person C knows	<p>The existence of persons A and B</p> <p><i>[Exposure_2 treatment only: You (person B) have identified the person A with whom you are matched but person A cannot identify you]</i></p> <p>The two possible messages</p> <p><i>[Chat treatment only: The chat opportunity between you and person A]</i></p>

<sup>16</sup> This sentence was not displayed in the Snitch treatment.

<sup>17</sup> The sentence "you do not make a decision" was not displayed in the Snitch treatment.

The message that person A decided to send to him  
*[Snitch treatment only: The fact that receiving a message from you means that person A's message is not true  
The fact that person A does not know whether or not person C received a message from you.]*

You (person B) know the message sent by person A to person C.  
His payoff after choosing an option

Note: C knows neither the payoffs at stake, nor the payoffs associated with each option, nor which option is the best for each of the three persons

*[Chat treatment only: nor the content of the discussion between you and person A.]*

All decisions are taken only once.

At the end of the session, you, person A and person C will only be informed about one's own payoff. In the second part, you will no longer be matched with any participant.

Please, read these instructions again. If you have any question, raise your hand and we will answer your questions in private.

## Part 2

*[The instructions for Part 2 are similar to those of the senders.]*

### C) Instructions for the receivers

Thank you for participating in this short experimental session on decision-making. This session consists of two completely independent parts. We have distributed the instructions for the first part. You will receive the instructions for the last part later.

You will be paid individually in cash in a separate room by someone who is not aware of the content of the experiment.

For the duration of the experiment, your decisions are anonymous; you will never enter your name into the computer. Please remain silent throughout the experiment. These instructions are for your own use only.

## Part 1

You are called "**person C**" and you will be randomly matched with two other people who participate in this session, called "person A" and "person B."

You will never know the identity of the other two participants and the other participants will never know your identity.

*[In the Exposure\_2 treatment: In contrast, the person B with whom you are matched will be able to identify the person A with whom you are matched in the room. Indeed, at the beginning of this part before decisions are made, we will ask each person A to stand up, one by one. We will call each person A by the number on his tag and person A will have to stand up (for example, if person A has received the tag with number 1, it means that this person is person A number 1). The person B with whom you are matched knows the number of the tag of person A and thus this person will be able to identify person A in the room. In contrast, person A will not be able to identify person B with whom you are matched. We asked person B not to make any sign that would allow person A to identify him in the room. ]*

Two possible payment options, called "option X" and "option Y", are available to you and your co-participants. Each option specifies a payoff for person A, a payoff for person B and a payoff for person C (you).

Payoffs depend on the payment option that you will select, X or Y. We have shown the two payment options to person A and to person B. The only information that you have available to make your choice is the message sent to you by person A.

The two possible messages are the following:

Message 1: "Option X earns you more money than option Y."

Message 2: "Option Y earns you more money than option X."

You will be informed about the message sent to you by person A. Then, we will ask you to choose between option X and option Y. Your choice will determine the earnings of the three persons. You will never know the amounts that were offered in the option that you did not select (that is to say, you will never know if the message of person A was true or not).<sup>18</sup> In addition, you will never know the amounts that person A and person B earn depending on the different options.

Person B receives the same information as person A: B is informed of the payoffs associated with each option, X and Y, and knows the two possible messages person A can send to you. Person B can observe the message A has chosen to send to you and person A knows this. Person B makes no decision.<sup>19</sup>

*[Chat treatment only: Before person A makes his decision about the message to send to you, person A has the opportunity to interact with person B using a chat box that appears on their screen. The only restriction regarding their interaction is that it is forbidden to identify oneself or ask the other person for identification, use foul language or threats. Persons A and B have 3 minutes to chat. Persons A and B know that you are aware of their opportunity to chat but that you will never know the content of their discussion. Once the 3 minutes have elapsed, only person A will enter on his computer the message that he has decided to send to you. Person B is informed about the message that person A sends to you but cannot send you any message himself.]*

*[Snitch treatment only: Before seeing the message that person A has selected, person B must decide whether or not to send you the following message in the case the message sent by person A is not true: The message from person A is not true. Option ... earns you more money than option ...."]*

*- If person A's message is not true, person B's message will be sent to you only if person B has decided to send you a message.*

*- If person A's message is true, person B cannot send a message to you.*

*Person B will then be informed of the message that person A actually chose to send to you. But person A will not know whether person B has decided to send a message to you if his message is not true. If you receive a message from person B, it means that the message of person A is not true. If you do not receive a message from person B, you do not know whether this is because person A's message is true or because person B has decided not to send you a message.]*

The following table summarizes the information possessed by persons A, B, and C (you):

Person	Information
What person A knows	<p>The existence of persons B and C</p> <p><i>[Exposure 2 treatment only: The fact that person B has identified him in the room]</i></p> <p>The payoffs of each person with each option</p> <p>Which option is denoted "X" and which is denoted "Y"</p> <p>The two possible messages</p> <p><i>[Chat treatment only: The possibility to chat with person B]</i></p> <p><i>[Snitch treatment only: The possibility for person B to send the other message to you if person A's message is not true. Note: If his message is not true, person A does not know if person B has sent or not a message to you.]</i></p> <p>Person B knows the message he sent to you.</p> <p>His payoff after you have chosen an option</p>
What person B knows	<p>The existence of persons A and C</p> <p><i>[Exposure 2 treatment only: person B can identify person A with whom he is matched]</i></p> <p>The earnings of each person with each option</p> <p>Which option is denoted "X" and which is denoted "Y"</p> <p>The two possible messages</p> <p><i>[Chat treatment only: The possibility to chat with person A]</i></p>

<sup>18</sup> The sentence: "that is to say, you will never know if the message of person A was true or not" was not displayed in the Snitch treatment.

<sup>19</sup> This sentence was not displayed in the Snitch treatment.

	<i>[Snitch treatment only: The possibility to send the other message to you only if person A's message is not true The fact that person A does not know whether or not person B sent a message to you.]</i> The message person A sent to you His payoff after you have chosen an option
You	The existence of persons A and B <i>[Exposure_2 treatment only: person B has identified the person A with whom he is matched in the room but person A cannot identify person B]</i> The two possible messages <i>[Chat treatment only: The chat opportunity between person A and person B]</i> <i>[Snitch treatment only: The fact that receiving a message from person B means that person A's message is not true The fact that person A does not know whether or not person C received a message from you.]</i> The message sent by person A to you Your payoff after choosing an option  Note: you know neither the payoffs at stake, nor the payoffs associated with each option, nor which option is the best for each of the three persons <i>[Chat treatment only: nor the content of the discussion between you and person A.]</i>

All decisions are taken only once.

At the end of the session, you, person A and person B will only be informed about one's own payoff. In the second part, you will no longer be matched with any participant.

Please, read these instructions again. If you have any question, raise your hand and we will answer your questions in private.

## Part 2

*[The instructions for Part 2 are similar to those of the senders. ]*

## Online Appendix 2. Instructions for the ten-option treatments (translated from French)

*The instructions are for the two ten-option treatments. The parts of the instructions that are specific to the Exposure\_10\_Conflict treatment are indicated into brackets in italics.*

### A) Instructions for the senders

Thank you for participating in this short experimental session on decision-making. This session consists of two completely independent parts. We have distributed the instructions for the first part. You will receive the instructions for the last part later.

You will be paid individually in cash in a separate room by someone who is not aware of the content of the experiment.

For the duration of the experiment, your decisions are anonymous; you will never enter your name into the computer. Please remain silent throughout the experiment. These instructions are for your own use only.

### Part 1

You are called "**person A**" and you will be randomly matched with two other people who participate in this session, called "person B" and "person C."

You will never know the identity of the other two participants. Person C with whom you are matched will never know your identity or that of person B. In contrast, person B with whom you are matched will be able to identify you in the room. Indeed, at the beginning of this part before you make your decision, we will ask each person A to stand up, one by one. We will call you by the number on your tag and you will have to stand up (for example, if you have received the tag with number 1, it means that you are person A number 1). Person B with whom you are matched knows the number of your tag, so that this person will be able to identify you in the room. In contrast, you will not be able to identify the person B you are matched with. We asked person B not to make any sign that would allow you to identify him in the room. Person C knows that person B has identified you and that you cannot identify person B.

Ten possible payment options (numbered from 1 to 10) are available to you and your co-participants in this part. Before the beginning of the experiment, the computer program has randomly selected one number between 1 and 10. This number is the number of the option that pays the following amounts:

€ 5 for you, 6 € for person B and 15 € for person C

The nine other options pay the following amounts:  
€ 15 for you, 7 € for person B and 4 € for person C.

*[In the Exposure\_10\_Conflict treatment, the previous allocations are replaced by the following paragraphs: This number is the number of the option that pays the following amounts:*

*€ 5 for you, 15 € for person B and 6 € for person C*

*The nine other options pay the following amounts:*

*€ 15 for you, 7 € for person B and 4 € for person C]*

At the beginning of this part, you and person B (but not person C) learn which number has been randomly drawn.

Person C must choose between the ten options to determine which option will be implemented. The only information that person C possesses is the information that you send him in a message. Person C will never know the monetary payoffs associated with each option, nor the differences in earnings. Person C will never be informed of which option is best for each of the three persons.

We will ask you to choose one of the ten following messages you send to person C:

Message 1: "Option 1 earns you more money than the other options."

Message 2: "Option 2 earns you more money than the other options."

Message 3: "Option 3 earns you more money than the other options."

Message 4: "Option 4 earns you more money than the other options."  
 Message 5: "Option 5 earns you more money than the other options."  
 Message 6: "Option 6 earns you more money than the other options."  
 Message 7: "Option 7 earns you more money than the other options."  
 Message 8: "Option 8 earns you more money than the other options."  
 Message 9: "Option 9 earns you more money than the other options."  
 Message 10: "Option 10 earns you more money than the other options."

We will forward your message to person C and ask him to choose between the ten options. The choice of person C will determine the earnings of the three persons. However, person C will never know the amounts that were offered in the options that were not selected (that is to say, he will never know if your message was true or not). In addition, person C will never know the amounts that you and person B earn depending on the different options.

Person B receives the same information as you: B is informed of the payoffs associated with each option, and knows the ten messages that you can send to person C. Person B is informed of the message you send to person C but he cannot send any message himself to person C. Person B makes no decision. Both person B and person C know that you know the payoffs associated with each option.

The following table summarizes the information possessed by persons A (you), B and C:

Person	Information
You	The existence of persons B and C The fact that the person B with whom you are matched has identified you in the room The payoffs of each person with each option The ten possible messages Person B knows the message you decided to send to person C. Your payoff after person C has chosen an option
What person B knows	The existence of persons A and C Person B with whom you are matched has identified you in the room The earnings of each person with each option The ten possible messages The message you (person A) decided to send to person C His payoff after person C has chosen an option
What person C knows	The existence of persons A and B Person B with whom you (person A) are matched has identified you in the room but you cannot identify who is person B The ten possible messages The message that you (person A) send to him Person B knows the message that you (person A) decided to send to person C. His payoff after choosing an option  Note: C knows neither the payoffs at stake, nor the payoffs associated with each of the ten options, nor which option is the best for each of the three persons

All decisions are taken only once.

At the end of the session, you, person B and person C will only be informed about one's own payoff. In the second part, you will no longer be matched with any participant.

Please, read these instructions again. If you have any question, raise your hand and we will answer your questions in private.

## Part 2

*[The instructions for Part 2 are similar to those of the other treatments. See Appendix 1.]*

-----

## B) Instructions for the observers

Thank you for participating in this short experimental session on decision-making. This session consists of two completely independent parts. We have distributed the instructions for the first part. You will receive the instructions for the last part later.

You will be paid individually in cash in a separate room by someone who is not aware of the content of the experiment.

For the duration of the experiment, your decisions are anonymous; you will never enter your name into the computer. Please remain silent throughout the experiment. These instructions are for your own use only.

### Part 1

You are called "**person B**" and you will be randomly matched with two other people who participate in this session, called "person A" and "person C."

The persons A and C with whom you are matched will never know your identity. You will never know the identity of person C. In contrast, you will be able to identify the person A with whom you are matched in the room. Indeed, at the beginning of this part before decisions are made, we will ask each person A to stand up, one by one. We will call each person A by the number on his tag and person A will have to stand up. You know the number of the tag of the person A with whom you have been matched: it is the number indicated on the tag we gave you (for example, if you have received the tag with number 1, it means that you are matched with person A number 1). Thus, you will be able to identify the person A with whom you are matched. In contrast, person A will not be able to identify you. We ask you not to make any sign that would allow person A to identify you in the room. Person C knows that you can identify person A and that person A cannot identify you.

Ten possible payment options (numbered from 1 to 10) are available to you and your co-participants in this part. Before the beginning of the experiment, the computer program has randomly selected one number between 1 and 10. This number is the number of the option that pays the following amounts:

€ 5 for person A, 6 € for you and 15 € for person C

The nine other options pay the following amounts:  
€ 15 for person A, 7 € for you and 4 € for person C.

*[In the Exposure\_10\_Conflict treatment, the previous allocations are replaced by the following paragraphs: This number is the number of the option that pays the following amounts:*

*€ 5 for you, 15 € for person B and 6 € for person C*

*The nine other options pay the following amounts:*

*€ 15 for you, 7 € for person B and 4 € for person C]*

At the beginning of this part, you and person A (but not person C) learn which number has been randomly drawn.

Person C must choose between the ten options to determine which option will be implemented. The only information that person C possesses is the information that person A send him in a message. Person C will never know the monetary payoffs associated with each option, nor the earnings gap. Person C will never be informed about which option is best for each of the three persons.

We will ask person A to choose one of the ten following messages to send to person C:

Message 1: "Option 1 earns you more money than the other options."

Message 2: "Option 2 earns you more money than the other options."

Message 3: "Option 3 earns you more money than the other options."

Message 4: "Option 4 earns you more money than the other options."

Message 5: "Option 5 earns you more money than the other options."

Message 6: "Option 6 earns you more money than the other options."

Message 7: "Option 7 earns you more money than the other options."

Message 8: "Option 8 earns you more money than the other options."  
 Message 9: "Option 9 earns you more money than the other options."  
 Message 10: "Option 10 earns you more money than the other options."

We will forward the message of person A to person C and ask him to choose between the ten options. The choice of person C will determine the earnings of the three persons. However, person C will never know the amounts that were offered in the options that were not selected (that is to say, he will never know if the message of person A was true or not). In addition, person C will never know the amounts that you and person A earn depending on the different options.

You are informed of the message that person A sends to person C. You have no decision to make. Both person A and person C know that: you know the payoffs associated with each option, the ten possible messages that can be sent by A, the message chosen by A; you do not make a decision.

The following table summarizes the information possessed by persons A, B (you), and C:

Person	Information
What person A knows	The existence of persons B and C The fact that you (person B) have identified him in the room The payoffs of each person with each option The ten possible messages You (person B) know the message he decided to send to person C. His payoff after person C has chosen an option
You	The existence of persons A and C You can identify person A with whom you are matched in the room The earnings of each person with each option The ten possible messages The message person A decided to send to person C Your payoff after person C has chosen an option
What person C knows	The existence of persons A and B You (person B) have identified the person A with whom you are matched but person A cannot identify you The ten possible messages The message that person A decided to send to him You (person B) know the message sent by person A to person C. His payoff after choosing an option  Note: C knows neither the payoffs at stake, nor the payoffs associated with each option, nor which option is the best for each of the three persons

All decisions are taken only once.

At the end of the session, you, person A and person C will only be informed about one's own payoff. In the second part, you will no longer be matched with any participant.

Please, read these instructions again. If you have any question, raise your hand and we will answer your questions in private.

## Part 2

*[The instructions for Part 2 are similar to those of the senders.]*

-----

### C) Instructions for the receivers

Thank you for participating in this short experimental session on decision-making. This session consists of two completely independent parts. We have distributed the instructions for the first part. You will receive the instructions for the last part later.

You will be paid individually in cash in a separate room by someone who is not aware of the content of the experiment.



For the duration of the experiment, your decisions are anonymous; you will never enter your name into the computer. Please remain silent throughout the experiment. These instructions are for your own use only.

## Part 1

You are called "**person C**" and you will be randomly matched with two other people who participate in this session, called "person A" and "person B."

You will never know the identity of the other two participants and the other participants will never know your identity.

In contrast, the person B with whom you are matched will be able to identify the person A with whom you are matched in the room. Indeed, at the beginning of this part before decisions are made, we will ask each person A to stand up, one by one. We will call each person A by the number on his tag and person A will have to stand up (for example, if person A has received the tag with number 1, it means that this person is person A number 1). The person B with whom you are matched knows the number of the tag of person A and thus this person will be able to identify person A in the room. In contrast, person A will not be able to identify person B with whom you are matched. We asked person B not to make any sign that would allow person A to identify him in the room.

Ten possible payment options (numbered from 1 to 10) are available for you and your co-participants in this part. Each option specifies a payoff for person A, a payoff for person B and a payoff for person C (you).

Payoffs depend on the payment option that you will select, X or Y. We have shown the ten payment options to person A and to person B. The only information that you have to make your choice is the message sent to you by person A.

The ten possible messages are the following:

Message 1: "Option 1 earns you more money than the other options."  
 Message 2: "Option 2 earns you more money than the other options."  
 Message 3: "Option 3 earns you more money than the other options."  
 Message 4: "Option 4 earns you more money than the other options."  
 Message 5: "Option 5 earns you more money than the other options."  
 Message 6: "Option 6 earns you more money than the other options."  
 Message 7: "Option 7 earns you more money than the other options."  
 Message 8: "Option 8 earns you more money than the other options."  
 Message 9: "Option 9 earns you more money than the other options."  
 Message 10: "Option 10 earns you more money than the other options."

You will be informed of the message sent to you by person A. Then, we will ask you to choose between the ten options by choosing a number between 1 and 10. Your choice will determine the earnings of the three persons. You will never know the amounts that were offered in the options that you did not select (that is to say, you will never know if the message of person A was true or not). In addition, you will never know the amounts that person A and person B earn depending on the different options.

Person B receives the same information as person A: B is informed of the payoffs associated with each of the ten options, and knows the ten possible messages person A can send to you. Person B can observe the message A has chosen to send you and person A knows it. Person B makes no decision.

The following table summarizes the information possessed by persons A, B, and C (you):

Person	Information
What person A knows	The existence of persons B and C The fact that person B has identified him in the room The payoffs of each person with each option The ten possible messages Person B knows the message person A decided to send to you. His payoff after you have chosen an option
What person B knows	The existence of persons A and C Person B can identify person A with whom he is matched The earnings of each person with each option

	The ten possible messages The message person A decided to send to you His payoff after you have chosen an option
You	The existence of persons A and B Person B has identified the person A with whom he is matched in the room but person A cannot identify person B The ten possible messages The message sent by person A to you Your payoff after choosing an option  Note: you know neither the payoffs at stake, nor the payoffs associated with each option, nor which option is the best for each of the three persons

All decisions are taken only once.

At the end of the session, you, person A and person B will only be informed about one's own payoff. In the second part, you will no longer be matched with any participant.

Please, read these instructions again. If you have any question, raise your hand and we will answer your questions in private.

## Part 2

*[The instructions for Part 2 are similar to those of the senders.]*

**Online Appendix Table A1: Determinants of the senders' lying rate**

	(1)	(2)	(3)	(4)	(5)
	All	Lying is disadvantageous	Lying is advantageous	Lying is disadvantageous	Lying is advantageous
Exposure treatment	-0.016 (0.093)	0.003 (0.125)	0.012 (0.096)	-0.018 (0.127)	0.031 (0.117)
Chat treatment	-0.015 (0.071)	-0.129 (0.100)	0.043 (0.115)	-0.134 (0.103)	0.051 (0.134)
Snitch treatment	0.083 (0.095)	0.352** (0.144)	-0.017 (0.167)	0.305** (0.127)	-0.011 (0.157)
Sender's belief	0.007*** (0.002)	-0.005 (0.005)	0.004 (0.004)	-0.003 (0.004)	0.004 (0.003)
Female	-	-	-	-0.075 (0.120)	-0.003 (0.167)
Age (in years)	-	-	-	-0.009 (0.014)	-0.007 (0.009)
Machiavellean score (60-140)	-	-	-	-0.004 (0.007)	0.013 (0.009)
Risk-aversion (0-100)	-	-	-	0.003 (0.002)	-0.001 (0.003)
Constant	0.103 (0.126)	0.414** (0.178)	0.375 (0.246)	0.881 (0.967)	-0.670 (0.801)
Number of observations	129	65	64	65	64
R-squared	0.092	0.126	0.011	0.178	0.070

Notes: Sender's belief is the reported belief about the percentage of receivers that will follow their message. Robust standard errors in parentheses, clustered at the session level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$